

A Neo-Piagetian Framework for Understanding Anthropomorphic Response Patterns in Conversational Agents

San Hong
Department of Industrial
Engineering
Seoul National University
Seoul, Korea
saanhong@snu.ac.kr

Yejin Kang
Department of Industrial
Engineering
Seoul National University
Seoul, Korea
yejinjasmine@snu.ac.kr

Woojin Park
Department of Industrial
Engineering
Seoul National University
Seoul, Korea
woojinpark@snu.ac.kr

ABSTRACT

The rapid proliferation of conversational agents powered by large language models presents a fundamental challenge to human-AI interaction: users systematically vary in how they perceive and interact with these systems, yet no coherent theoretical framework explains these patterns. This paper proposes a Neo-Piagetian developmental framework to understand anthropomorphic response patterns in conversational agent interactions. Drawing on domain-specific cognitive development theory, we demonstrate that users progress through distinct developmental stages—preoperational, concrete operational, and formal operational—characterized by qualitatively different ways of understanding AI systems. Through conceptual framework methodology, we synthesize empirical evidence showing how users with low AI literacy exhibit animistic thinking and egocentrism (preoperational), while those with moderate literacy develop functional classifications without grasping underlying mechanisms (concrete operational), and high-literacy users achieve abstract reasoning about computational architectures (formal operational). We introduce a cyclical model of AI literacy development driven by Piagetian mechanisms of assimilation, disequilibrium, and equilibration, where unexpected AI behaviors trigger cognitive conflict that propels developmental progression. This framework resolves contradictions in existing literature by showing that anthropomorphic responses reflect users' developmental stage rather than random variation, and that the same individual can exhibit different cognitive levels across domains. The framework has significant implications for AI design, deployment, and education, suggesting that systems should adapt to users' developmental stages and that literacy interventions can accelerate progression beyond primitive anthropomorphic thinking. As conversational agents become increasingly sophisticated in mimicking human communication, understanding these developmental dynamics becomes critical for designing systems that balance accessibility with appropriate user understanding.

KEYWORDS

AI literacy, Human-AI interaction, Cognitive development, Chatbot interaction

1 INTRODUCTION

The widespread adoption of conversational agents has created a significant challenge: while we understand that users respond to AI with varying degrees of anthropomorphism, we lack a coherent explanation for why these responses differ so systematically across individuals and contexts. These systematic differences manifest in predictable patterns. Users consistently cluster into distinct interaction styles, from those who engage in extensive social dialogue to those maintaining purely functional exchanges [1]. These patterns correlate strongly with AI literacy levels—users with limited AI understanding show markedly different response patterns, particularly benefiting from rapport-building strategies that more experienced users find unnecessary [2]. Beyond individual differences, contextual variations follow equally systematic patterns. Physical embodiment dramatically amplifies anthropomorphic responses, with robots generating stronger reactions than chatbots, particularly in emotionally-charged domains like healthcare and education [3]. Roesler et al.'s meta-analysis discovered a consistent pattern: anthropomorphic design succeeds in social contexts but fails in industrial applications [4]. Without understanding the cognitive mechanisms underlying these systematic variations, we cannot design AI systems that appropriately adapt to diverse user needs or develop effective AI literacy interventions.

Current research suggests fundamental cognitive phenomena that demand theoretical explanation. Users do not simply project human qualities onto machines based on surface features; rather, they engage in complex cognitive processes involving intentionality attribution and social presence perception [5]. This cognitive foundation aligns with the Computers Are Social Actors (CASA) paradigm [6], yet extends beyond it—users appear to be at different cognitive stages in their ability to understand and interact with AI systems. The literacy-response relationship observed in recent studies hints at deeper cognitive processes at work, suggesting that users' understanding of AI may develop through identifiable stages rather than accumulating linearly.

We propose that Neo-Piagetian cognitive development theory may offer the missing theoretical framework. Unlike traditional Piagetian approaches that assume uniform cognitive development across domains, Neo-Piagetian theory recognizes that adults can

exhibit varying developmental levels in different knowledge areas. This domain-specificity could explain a crucial observation: the same individual who demonstrates sophisticated formal reasoning in familiar domains may revert to animistic, preoperational thinking when encountering AI systems. The key idea is that AI literacy might represent a distinct cognitive domain where users progress through identifiable developmental stages.

However, whether Neo-Piagetian theory actually applies to human-AI interaction remains unexplored. While the patterns we observe—from animistic attribution to systematic understanding—suggest developmental progression, no existing research has examined whether these patterns truly reflect Neo-Piagetian stages. The systematic variations in anthropomorphism and the domain-specific nature of AI understanding provide compelling foundations for applying this theoretical lens, but empirical investigation is needed to determine its validity.

This paper addresses the research question: Do conversational agent users' anthropomorphic responses and interaction patterns reflect identifiable Neo-Piagetian developmental stages, and if so, how? We first examine whether user behaviors can be meaningfully mapped onto cognitive developmental stages, then characterize these stages if they exist. By potentially establishing this theoretical connection, we aim to provide both theoretical coherence for disparate empirical findings and practical guidance for adaptive AI design.

2 THEORETICAL BACKGROUND

2.1 From Piaget to Neo-Piagetian: The Domain-Specificity Revolution

Traditional Piagetian theory posited universal cognitive stages progressing from sensorimotor through preoperational and concrete operational to formal operational thinking [7, 8]. Central to Piaget's framework was the assumption of domain-generality—cognitive structures developed holistically, affecting all areas of reasoning simultaneously. A child reaching formal operations would theoretically demonstrate abstract thinking across mathematics, social reasoning, and scientific understanding equally [9].

However, empirical evidence increasingly challenged this domain-general assumption. Children and adults demonstrated uneven cognitive profiles, excelling in certain domains while struggling in others [10, 11]. This led to the Neo-Piagetian revolution, which retained Piaget's developmental progression while introducing domain-specificity—the recognition that cognitive development occurs independently across different knowledge areas [12, 13].

Fischer's Dynamic Skills Theory reconceptualized cognitive development as context-specific skill acquisition rather than abstract logical structures, suggesting varying developmental levels across knowledge domains [14]. Case's Central Conceptual Structures [12, 15] elaborated this through meaning networks within experiential domains like number, space, and social interaction. Demetriou's Specialized Structural Systems (SSS) [13] proposed multiple autonomous thinking systems for processing different reality aspects. The Social SSS, governing social relationship understanding, becomes particularly relevant for AI anthropomorphism—users with underdeveloped AI-specific Social SSS default to primitive cognitive patterns.

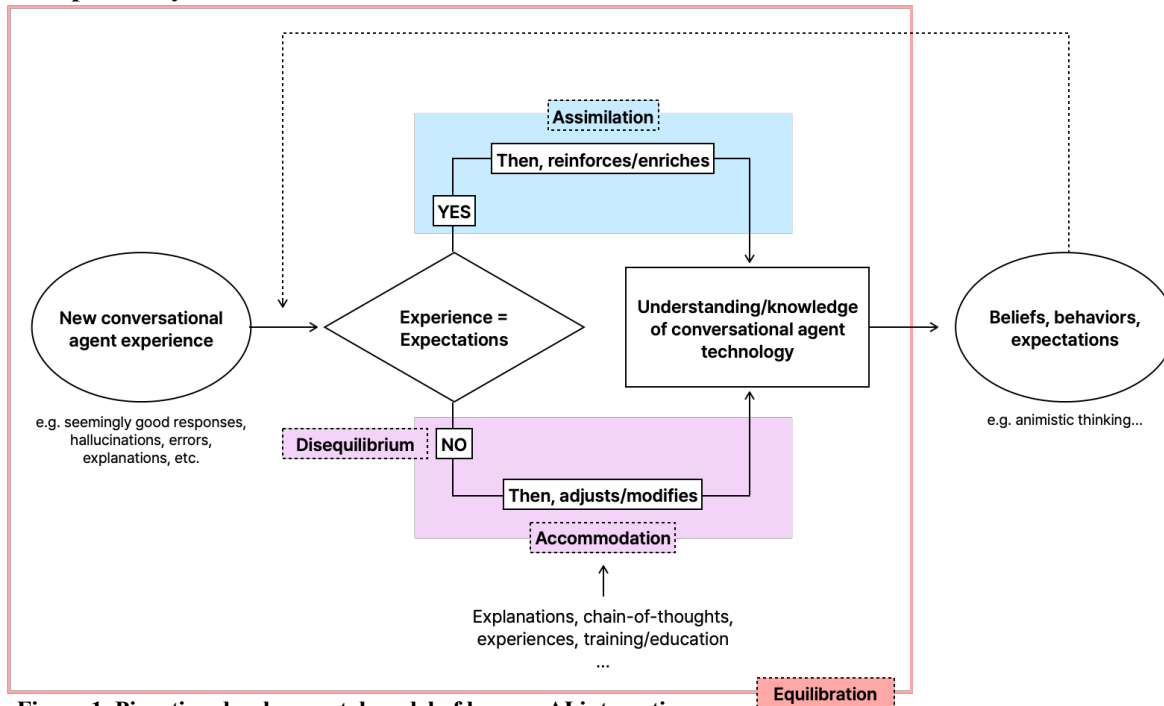


Figure 1. Piagetian developmental model of human-AI interaction

2.2 AI as a Novel Cognitive Domain

Conversational agents represent an unprecedented cognitive challenge that makes Neo-Piagetian analysis particularly appropriate and illuminating. Unlike traditional domains grounded in evolutionary preparedness [16] or cultural scaffolding, they constitute a genuinely novel domain requiring new cognitive structures. Three factors illustrate why Neo-Piagetian interpretation is especially promising for understanding conversational agent interactions.

First, the domain-specificity principle explains why accomplished professionals can exhibit primitive thinking patterns when encountering AI. A physicist demonstrating formal operations in their field may resort to animistic, preoperational thinking when interpreting chatbot responses—their AI-specific cognitive structures remain underdeveloped despite expertise elsewhere [17, 18].

Second, the rapid technological emergence means users lack established cultural frameworks for understanding AI behavior. Without inherited wisdom or intuitive models [19], individuals must construct their understanding from scratch, recapitulating developmental stages within this specific domain. Users with underdeveloped AI-specific Social SSS default to primitive cognitive patterns, treating mechanical responses as intentional communication [20, 21].

Third, AI's unique ontological status—neither clearly animate nor inanimate—creates cognitive ambiguity that reveals developmental stages particularly clearly [22, 23]. Users must actively construct mental models to interpret agent behaviors, and these constructions follow predictable developmental patterns from anthropomorphic projection to systematic understanding of computational processes [23, 24].

3 METHODS

We employed Jabareen's conceptual framework methodology [25] to synthesize multidisciplinary literature into a coherent theoretical framework. Through systematic database searches, we identified 38 papers examining human interactions with conversational agents (chatbots, virtual assistants, and LLM-based systems) that included empirical data on user-agent interactions with detailed behavioral observations.

From each paper, we extracted reported instances of users' anthropomorphic responses and interaction patterns, along with contextual information including user AI literacy levels, situational contexts, and specific agent types. We used an inductive approach to identify patterns across these cases, discovering clusters of similar behaviors that suggested distinct developmental stages.

These emergent patterns suggest strong parallels with classical Piagetian stages (preoperational, concrete operational, and formal operational thinking). Based on this alignment, we developed a cyclical model incorporating the Piagetian mechanisms of assimilation, accommodation, and equilibration to explain transitions between stages. For each stage, we synthesized common characteristics across cases to create comprehensive descriptions

grounded in the empirically observed behaviors documented in existing research.

4 RESULTS

A Neo-Piagetian Developmental Framework for AI Anthropomorphism. Our analysis of behavioral instances surfaced three distinct developmental stages that suggest Piagetian cognitive stages. As shown in Figure 1, the developmental dynamics of AI literacy can be understood through the Piagetian mechanisms of assimilation, accommodation, and equilibration. Figure 1 illustrates how users develop their understanding of conversational agents through a cyclical process. Starting from the left with "New conversational agent experiences," users encounter AI systems with their existing mental models. If the agent's behavior matches expectations (YES path), assimilation occurs (blue box, top), where users integrate the experience into their current understanding. If the behavior violates expectations (NO path), users experience disequilibrium—a cognitive conflict between experience and expectation. This triggers accommodation (pink box, bottom), where users must adjust their mental models to resolve the conflict. The process results in updated beliefs, behaviors, and expectations (right side). The arrow returning to the left shows the cyclical nature: users re-enter the cycle with their updated understanding, potentially triggering new rounds of assimilation or accommodation. Through repeated cycles, users progress from preoperational to concrete operational to formal operational thinking about AI.

Assimilation refers to the process by which users fit new encounters with conversational agents into their existing mental models—for example, interpreting fluent responses as evidence of human-like understanding or applying prior categories such as "chatbot" or "search engine." When the agent produces unexpected behaviors—such as hallucinations, repetitive answers, or logically inconsistent outputs—these events trigger disequilibrium, a state of cognitive conflict in which prior schemas no longer adequately explain system behavior. In the AI domain, disequilibrium manifests as confusion, mistrust, or surprise when the system violates assumptions about coherence or reliability [26, 27].

Accommodation occurs as users attempt to resolve this conflict by actively restructuring their mental models—adjusting prior assumptions, incorporating new explanatory frameworks, or adopting alternative strategies for interpreting agent behavior. Existing research documents that users who experienced repeated disequilibrium events gradually developed more sophisticated mental models, progressing from anthropomorphic to mechanistic understanding. Through this process, perceptions are reformed and eventually reach a new equilibrium, where users adjust their literacy and anthropomorphic attributions in light of both the system's limitations and its affordances [28].

Building on this cyclical perspective, the behavioral patterns we discovered mapped onto Piaget's classic stages of cognitive development. This framework enables us to conceptualize how individuals with varying levels of AI literacy anthropomorphize,

evaluate, and ultimately reformulate their understanding of conversational agents. The following sections describe the characteristics we extracted from behavioral instances within each developmental stage.

Preoperational Stage (Low AI Literacy): Analysis of behavioral instances from users with low AI literacy revealed consistent patterns of animistic thinking. Users attributed consciousness and emotions to agents, with common behaviors including: apologizing to chatbots for "bothering" them, expressing concern about "hurting" the AI's feelings, and believing agents could feel tired or frustrated. Egocentrism manifested as assuming shared emotional states and contextual understanding. The Eliza effect [29] was evident in multiple instances where users interpreted mechanical language repetition as authentic connection.

Characteristic features extracted from this stage included: centration on social cues (friendly language, emoticons) while overlooking mechanical indicators (repetitive responses, logical inconsistencies). As Peter et al. [30] note, LLM-based systems can convincingly mimic empathy, persuasion, and role-play, creating what they call "anthropomorphic seduction," where users mistake fluent interaction for genuine understanding. Sarkar et al. [31] further highlight that in mental health applications, such dynamics reproduce the Eliza effect at scale.

Concrete Operational Stage (Moderate AI Literacy): Users in this stage demonstrated rule-based thinking about conversational agents. Common behavioral patterns included: testing agents with similar questions to verify consistency, categorizing different AI systems by function (search/translation/recommendation), and developing practical heuristics ("AI is good at facts but bad at opinions"). These users showed conservation concepts specific to AI systems, understanding that agents operate consistently regardless of time or user identity.

Literature reports that users at this stage could classify agents according to observable functions or access modalities (free vs. subscription-based) but lacked understanding of underlying mechanisms. Wienrich and Carolus [32] show that such functional competence with smart speakers does not necessarily translate into deeper conceptual understanding—ownership and frequent use did not increase conversational agent literacy scores, and misconceptions persisted.

Formal Operational Stage (High AI Literacy): Behavioral instances from highly literate users showed systematic hypothesis testing and abstract reasoning. Characteristic behaviors included: deliberately probing edge cases to understand system limitations, discussing probabilistic outputs and training data, and maintaining clear boundaries between human cognition and computational processes. These users demonstrated hypothetical-deductive reasoning about invisible architectures (neural networks, algorithms, databases).

Analysis of this stage indicates that users no longer relied on surface-level cues but instead interrogated underlying mechanisms through systematic testing. Liu et al. [33] emphasize that proactive agents designed with "inner thoughts" and reasoning chains invite such advanced interrogation, where high-literacy users can

differentiate between simulated cognition and genuine human thought. Desolda et al. [34] show how developers' mental models shape their ability to predict, validate, and customize AI behavior, underscoring that abstract reasoning is essential for aligning expectations with probabilistic outputs.

5 DISCUSSION

Our Neo-Piagetian framework transforms AI anthropomorphism from unexplained variance into predictable developmental patterns. By mapping user behaviors onto cognitive stages, we provide the theoretical mechanism missing from existing accounts.

The framework extends CASA theory [6, 35, 36] by explaining why social responses vary systematically—preoperational users exhibit heightened susceptibility to social cues through centration, while concrete and formal thinkers progressively decenter to examine functional patterns and computational mechanisms. Similarly, we operationalize Epley et al.'s [37] three-factor theory by mapping their drivers onto stage-specific operations: preoperational users' high accessibility of human schemas leads to animistic projection, concrete thinkers satisfy effectance needs through input-output pattern recognition, and formal reasoners maintain clear boundaries between human and computational processes through hypothetical-deductive testing.

This developmental perspective illuminates why device ownership alone fails to increase literacy [32]—exposure without developmental scaffolding cannot guarantee advancement from concrete to formal thinking. The framework also predicts context-dependent effects: Roesler et al.'s [4] meta-analysis showing anthropomorphism's success in social domains but failure in industrial settings reflects how preoperational users benefit from social scaffolding while formal reasoners prioritize transparency and predictability.

The assimilation-accommodation-equilibration cycle links micro-interactions to macro-development. When AI behaviors violate expectations, the resulting disequilibrium creates restructuring opportunities that differentially support stage progression. This mechanism aligns with dynamic CASA perspectives [39, 40] while specifying how interventions can facilitate development. By identifying developmental positions, we can predict anthropomorphic responses, design stage-appropriate interfaces, and scaffold AI literacy progression [38]. The framework thus bridges empirical observation and practical application, demonstrating how theoretical innovation guides effective human-AI interaction design and education.

ACKNOWLEDGMENTS

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No.RS-2022-II2209842)

REFERENCES

- [1] Galland, L., Pelachaud, C., & Pecune, F. (2022). Adapting conversational strategies in information-giving human-agent interaction. *Frontiers in Artificial Intelligence*, 5, 1029340.
- [2] Baihaqi, M. Y., Contreras, A. G., Kawano, S., & Yoshino, K. (2024). Rapport-driven virtual agent: Rapport building dialogue strategy for improving user experience at first meeting. arXiv preprint arXiv:2406.09839.
- [3] Blut, M., Wang, C., Wunderlich, N. V., & Brock, C. (2021). Understanding anthropomorphism in service provision: A meta-analysis of physical robots, chatbots, and other AI. *Journal of the Academy of Marketing Science*, 49(4), 632–658.
- [4] Roesler, E., Manzey, D., & Onnasch, L. (2021). A meta-analysis on the effectiveness of anthropomorphism in human-robot interaction. *Science Robotics*, 6(58), eabj5425. <https://doi.org/10.1126/scirobotics.abj5425>
- [5] Kim, J., & Im, I. (2023). Anthropomorphic response: Understanding interactions between humans and artificial intelligence agents. *Computers in Human Behavior*, 139, 107512.
- [6] Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, 56(1), 81–103.
- [7] Piaget, J., & Cook, M. (1952). *The origins of intelligence in children* (Vol. 8, No. 5, pp. 18–1952). New York: International Universities Press.
- [8] Piaget, J., & Inhelder, B. (1969). *The psychology of the child*. New York: Basic.
- [9] Lourenço, O., & Machado, A. (1996). In defense of Piaget's theory: A reply to 10 common criticisms. *Psychological Review*, 103(1), 143.
- [10] Higgins, E. T., Feldman, N. S., & Ruble, D. N. (1980). Accuracy and differentiation in social prediction: A developmental perspective. *Journal of Personality*, 48(4), 520–540.
- [11] Gardner, H., & Wolf, D. (1983). Waves and streams of symbolization: Notes on the development of symbolic capacities in young children. In *The acquisition of symbolic skills* (pp. 19–42). Springer.
- [12] Case, R. (1985). *Intellectual development: Birth to adulthood*. Academic Press.
- [13] Demetriou, A., Christou, C., Spanoudis, G., Platsidou, M., Fischer, K. W., & Dawson, T. L. (2002). The development of mental processing: Efficiency, working memory, and thinking. *Monographs of the Society for Research in Child Development*, i–167.
- [14] Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review*, 87(6), 477.
- [15] Case, R., Okamoto, Y., Griffin, S., McKeough, A., Bleiker, C., Henderson, B., & Stephenson, K. M. (1996). The role of central conceptual structures in the development of children's thought. *Monographs of the Society for Research in Child Development*, 61(1–2, Serial No. 246).
- [16] Cummins, D. D., & Cummins, R. (1999). Biological preparedness and evolutionary explanation. *Cognition*, 73(3), B37–B53. [https://doi.org/10.1016/s0010-0277\(99\)00062-1](https://doi.org/10.1016/s0010-0277(99)00062-1)
- [17] Keil, F. C. (2006). Explanation and understanding. *Annual Review of Psychology*, 57, 227–254. <https://doi.org/10.1146/annurev.psych.57.102904.190100>
- [18] Rozenblit, L., & Keil, F. (2002). The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive Science*, 26(5), 521–562. https://doi.org/10.1207/s15516709cog2605_1
- [19] Legare, C. H., & Nielsen, M. (2015). Imitation and innovation: The dual engines of cultural learning. *Trends in Cognitive Sciences*, 19(11), 688–699. <https://doi.org/10.1016/j.tics.2015.08.005>
- [20] Marchesi, S., Ghiglini, D., Ciardo, F., Perez-Osorio, J., Baykara, E., & Wykowska, A. (2019). Do we adopt the intentional stance toward humanoid robots? *Frontiers in Psychology*, 10, 450. <https://doi.org/10.3389/fpsyg.2019.00450>
- [21] Waytz, A., Heafner, J., & Epley, N. (2014). The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle. *Journal of Experimental Social Psychology*, 52, 113–117. <https://doi.org/10.1016/j.jesp.2014.01.005>
- [22] Smedegaard, C. V. (2022). Novelty knows no boundaries: Why a proper investigation of novelty effects within SHRI should begin by addressing the scientific plurality of the field. *Frontiers in Robotics and AI*, 9, 741478. <https://doi.org/10.3389/frobt.2022.741478>
- [23] Damiano, L., & Dumouchel, P. (2018). Anthropomorphism in human-robot co-evolution. *Frontiers in Psychology*, 9, 468. <https://doi.org/10.3389/fpsyg.2018.00468>
- [24] Gültekin, M. (2022). Human-social robot interaction, anthropomorphism and ontological boundary problem in education. *Psycho-Educational Research Reviews*, 11(3), 751–773. https://doi.org/10.52963/PERR_Biruni_V11.N3.11
- [25] Jabareen, Y. (2009). Building a conceptual framework: Philosophy, definitions, and procedure. *International Journal of Qualitative Methods*, 8(4), 49–62. <https://doi.org/10.1177/160940690900800406>
- [26] Kim, K., Hwang, J., Lee, Y., Lee, J., Kim, S.-E., & So, H.-J. (2025). Your thoughtful opponent: Embracing cognitive conflict with peer agent (Version 2). arXiv. <https://doi.org/10.48550/arXiv.2508.06955>
- [27] Massenon, R., Gambo, I., Khan, J. A., Agbonkese, C., & Alwadain, A. (2025). "My AI is lying to me": User-reported LLM hallucinations in AI mobile apps reviews. *Scientific Reports*, 15(1), 30397. <https://doi.org/10.1038/s41598-025-15416-8>
- [28] Haque, A. B., Islam, A. N., & Mikalef, P. (2023). Explainable artificial intelligence (XAI) from a user perspective: A synthesis of prior literature and problematizing avenues for future research. *Technological Forecasting and Social Change*, 186, 122120. <https://doi.org/10.1016/j.techfore.2022.122120>
- [29] Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36–45.
- [30] Peter, S., Riemer, K., & West, J. D. (2025). The benefits and dangers of anthropomorphic conversational agents. *Proceedings of the National Academy of Sciences*, 122(22). <https://doi.org/10.1073/pnas.2415898122>
- [31] Sarkar, S., Gaur, M., Chen, L. K., Garg, M., & Srivastava, B. (2023). A review of the explainability and safety of conversational agents for mental health to identify avenues for improvement. *Frontiers in Artificial Intelligence*, 6, 1229805. <https://doi.org/10.3389/frai.2023.1229805>
- [32] Wienrich, C., & Carolus, A. (2021). Development of an instrument to measure conceptualizations and competencies about conversational agents on the example of smart speakers. *Frontiers in Computer Science*, 3, 685277. <https://doi.org/10.3389/fcomp.2021.685277>
- [33] Liu, X. B., Fang, S., Shi, W., Wu, C., Igarashi, T., & Chen, X. ' (2025). Proactive conversational agents with inner thoughts. arXiv (Cornell University). <https://doi.org/10.1145/3706598.3713760>
- [34] Desolda, G., Eposito, A., Greco, F., Tucci, C., Buono, P., & Piccinno, A. (2025). Understanding user mental models in AI-driven code completion tools: Insights from an elicitation study. arXiv (Cornell University). <https://doi.org/10.48550/arxiv.2502.02194>
- [35] Nass, C., Steuer, J., & Tauber, E. R. (1994, April). Computers are social actors. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 72–78).
- [36] Lee, J. E. R., & Nass, C. I. (2010). Trust in computers: The computers-are-social-actors (CASA) paradigm and trustworthiness perception in human-computer communication. In *Trust and technology in a ubiquitous modern environment: Theoretical and methodological perspectives* (pp. 1–15). IGI Global.
- [37] Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, 114(4), 864–886. <https://doi.org/10.1037/0033-295X.114.4.864>
- [38] Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–16). <https://doi.org/10.1145/3313831.3376727>

- [39] Xu, K., Chen, X., & Huang, L. (2022). Deep mind in social responses to technologies: A new approach to explaining the Computers are Social Actors phenomena. *Computers in Human Behavior*, 134, 107321. <https://doi.org/10.1016/j.chb.2022.107321>
- [40] Gambino, A., Fox, J., & Ratan, R. (2020). Building a stronger CASA: Extending the Computers are Social Actors paradigm. *Human-Machine Communication*, 1, 71–86. <https://doi.org/10.30658/hmc.1.5>